DOCUMENT RESUME

ED 210 350 UD 021 785

AUTHOR Biernacki, Gerald J.; And Others
TITLE A Plan for Evaluation Services.
INSTITUTION Toledo Public Schools, Ohio.
SPONS AGENCY Ohio State Dept. of Education, Columbus.
PUB DATE 7 May 80
NOTE 45p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Change; *Formative Evaluation; Management by Objectives; *Models; Needs Assessment; *Program Costs; *Program Evaluation; Student Development; *Summative Evaluation
IDENTIFIERS Toledo Public Schools OH

ABSTRACT

The evaluation of instructional programs is necessary in order to document accomplishments, determine cost effectiveness, establish management control, and provide bases for improving the education of students. The responsibilities and technical requirements of an evaluation system involve the function of the following component areas: demographic information, program evaluation, program auditing, testing, and technical assistance. The evaluation process includes: (1) a generalized needs assessment and identification of the problem area; (2) a focused assessment involving program development, implementation and formative evaluation; and (3) a summative evaluation and dissemination of results. The organization and development of an evaluation system also involves the selection and training of staff as well as the implementation of a systemwide standardized testing program. (JCD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED210350

A PLAN FOR EVALUATION SERVICES

TOLEDO PUBLIC SCHOOLS

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Russell A. Working
Toledo Public Schools

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

UD021785

DR. DONALD J. STEELE, JR.
Superintendent

RUSSELL A. WORKING
Assistant Superintendent
Division of Planning and Evaluation

BOARD OF EDUCATION

GEORGE J. MUTTER
President
RICHARD C. DAOUST
Vice President
DR. SAMANTHA P. ADAMS
JUDE T. AUBRY
JUDY A. JONES

May 7, 1980

A PLAN FOR EVALUATION SERVICES

Prepared by

Gerald J. Biernacki, Executive Director, Research, Development and Evaluation
Russell A. Working, Assistant Superintendent Planning and Evaluation
Stephen Jurs, Professor, University of Toledo

Prepared under a grant from the State of Ohio Department of Education

May, 1980

ED021785

Section I

The General Framework for Evaluation

Evaluation needs to occur frequently in the administration of a school district. The effects of the instructional program and all the functions which support instruction require continuous evaluation both for documenting accomplishments (program benefits), determining cost effectiveness, establishing management control, and providing bases for improving the education of students. This paper defines and delineates the nature of evaluation that is conducted as a function of the Office of Research, Development and Evaluation, Division of Planning and Evaluation, Toledo Public Schools. A general framework for evaluation and its underlying assumptions are set forth and explained. Areas of evaluation are described in general and subsequent sections include detailed information about each evaluation component. These sections can be used as separate documents but they are included here to give a comprehensive description of the evaluation activities in the Toledo Public Schools. Supplemental information is included in the Appendix.

The Need for Evaluation

The school system exists to provide education through instructional programs and services for the children in the district. The reason for an evaluation component in the school system is to provide information that will support the instructional program. Much of the information is focused on assessing the quality of the instructional program. Testing of students and the evaluation of particular parts of the curriculum are major components of the evaluation effort which lead to improving the instructional programs.

Some of the evaluation activities are required by state and federal guidelines. Information is needed for many reports. Funds from these sources are usually linked to evaluation requirements. When the district receives

state or federal funds it is required to show how the money was spent and what benefits occurred as a result.

The range of essential and often mandated evaluation activities is extensive and is growing rather than lessening. School districts cannot meet this demand by just implementing a testing program. Evaluation offices have an expanding mission. This paper is an attempt to describe the component parts of the Evaluation Plans and to describe how the Office of Research, Development and Evaluation will carry out its mission.

Definition of Evaluation

The definition of evaluation that is appropriate for this Evaluation Plan is the following:

EVALUATION IS THE PROCESS OF COLLECTING, ANALYZING, AND REPORTING INFORMATION ABOUT SCHOOLS AND STUDENTS TO PROVIDE A BASIS FOR MAKING DECISIONS ABOUT PROGRAMS AND SERVICES.

This definition implies several important things. The evaluation must collect, analyze, and report information. Evaluators are to be included in deciding 1) what information should be gathered, 2) what is the best way to obtain that information, and 3) how the information should be analyzed so that there is clear and accurate communication of the results to the various individuals and groups who will act on the information. Another important implication of this definition is that the evaluator is providing information for decisions rather than actually making the decisions. This posture requires close cooperation between the evaluator and the decision makers so that there is assurance that the information which is gathered is actually the information that will be used in making decisions. (This definition is almost identical to that of Stufflebeam, et al. Educational Evaluation and Decision Making. Itasca: Peacock Publishers, 1971.)

Evaluations themselves may be designed for one or more purposes. Some

evaluations will merely monitor programs to attest that proposed activities actually occurred. Other evaluations will be focused on identifying successful and less successful aspects of the instructional program so that modifications can be made which will improve the programs. Cost effectiveness of programs will also be addressed within the above definition of evaluation.

The Bases for Decisions About Schools

The bases for decisions about schools include the actual performance of the students and the schools and the perceived performance which may or may not be consistent with the actual performance. Expectations, goals, and priorities for the schools come from both the perceived and actual performance of the schools. A major purpose of the evaluation is to provide accurate data on which to base expectations, goals, and priorities and also to report accurate information to the schools' various publics so that perceptions about the schools will be consistent with actual performance.

Figure 1 illustrates how this framework operates. In this figure the decisions about educational programs and priorities are based first on the perceived needs of schools and students. Over a period of time, the evaluation system provides a great deal of information from several sources.

The first source is an ongoing collection and reporting of demographic data about the population with particular attention on the Toledo school-age population. Projected enrollments and the socio-economic character of the community are examples of these data. The second source, testing, includes both the systemwide testing such as standardized achievement tests and special testing programs that individual schools may undertake. These latter tests might include measures of self concept or anxiety, for example, neither of which are part of the systemwide testing program. The third main information source comes from program evaluation. This component provides the accumulated results from the evaluations of a variety of educational programs such as Title

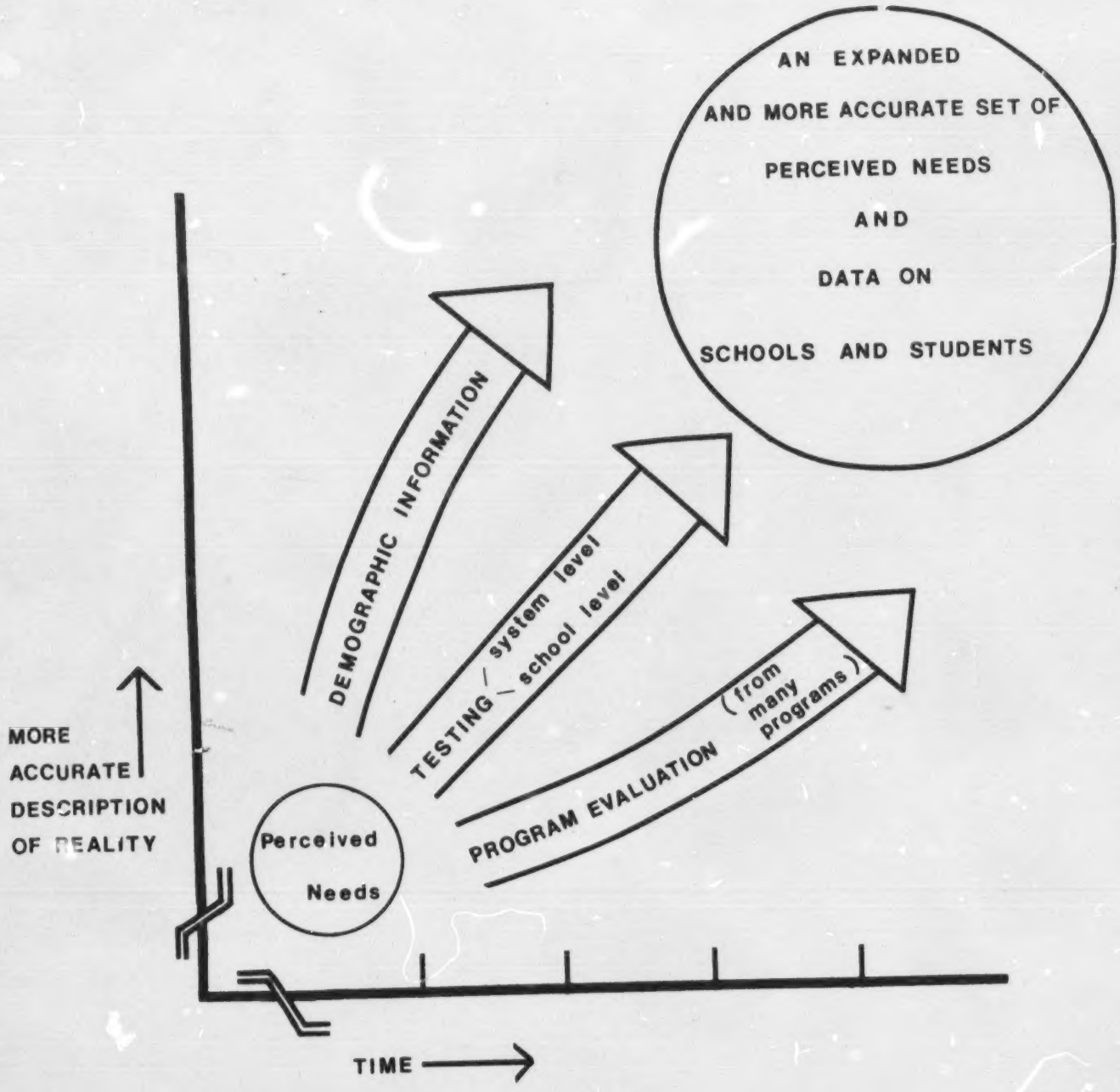


FIGURE 1. Evolving Bases for Decision Making

I Reading, English As a Second Language, etc. The results of each of these sets of studies provides additional information about the performance of students and schools.

The totality of this information leads to new and, it is hoped, more accurate sets of perceived needs of schools and students. Feedback of the data to the community should result in more accurate perceptions of the performance of the schools. The results of several years of programming for specific needs should result in rational, data-based decisions about those areas of need. Finally, longitudinally accumulated test data from all sources would also help to provide an accurate and valid basis for decisions about schools.

Assumptions On Which The Evaluation System is Based

The evaluation plan which follows is based upon the following set of guidelines.

Intended Components

1. An evaluation program must be more than a testing program although a testing program is an integral part of an evaluation program.

   This statement suggests that evaluation is to be a comprehensive program which is intended not just to gather data on a regular basis. It is intended to formulate investigation, synthesize information, and do other things in order to clarify successes, failures, cause and effect relationships between programs and results, community characteristics and test performance. All of these activities become a part of the evaluation program. It is not enough just to compile and report test results.

2. An evaluation program must provide both student oriented and program oriented data.

   Traditional offices of evaluation have acted as a repository of information about individual students. It is intended that the evaluation com-

ponent of the Toledo Public Schools also provides evaluative information about the various educational programs that occur within the schools. It is essential to have individual student data for purposes of aggregation or for identification of eligibility for certain programs. It is also essential to describe accurately the effects of various educational programs. The responsibility of the evaluation office include both of these functions.

3. An evaluation program must provide for the collection, analysis, and reporting of data in areas or categories which have been described by users as most helpful to them.

The evaluation component provides services to other users. This arrangement implies that it is necessary that there be close cooperation between the evaluation unit and those who use the evaluative information which is gathered. It is wasteful to gather data which is not useful to these groups. Similarly, it is dysfunctional to work independently of the potential users of the results of evaluations. Close cooperation is needed to best serve the interests and needs of the various user groups in the most efficient and productive manner.

4. An evaluation plan must be responsive to the information needs from a broad range of individuals -- teachers, parents, students, administrators, community groups, special interest groups, etc. -- in understandable language.

This point underscores the service orientation of the evaluation office by acknowledging the many audiences which the evaluation information can serve and by indicating that there is a responsibility to serve each of these groups rather than just one or two. The final phrase states that is the responsibility of the writers of evaluation results to make sure that the presentations communicate clearly to all groups of people that might use the results. Jargon, obscure statistics, and other confusing

statements should be avoided.

5. An evaluation plan must provide a level of flexibility which can accommodate (a) the collection, analysis, and reporting of new data sources, (b) the realignment of the "processes" of evaluation, and (c) the target audiences for which reports are prepared.

   Any evaluation system which does not have this desired level of flexibility will not be responsive to the evolving and expanding information needs of the potential users of evaluation results. A single evaluation process, no matter how general, may not serve all evaluation situations adequately.

6. An evaluation plan must focus on programs and services and therefore will not be used to evaluate individual teachers.

   Test scores will not be used to evaluate teachers. The teaching-learning process is quite complex. The quality of the teaching cannot be directly judged by the quality of learning. There are too many other variables which affect learning. The test scores of children are the results of the cumulative education of the children. The experiences the children have encountered conflict with the home environment and inner and external motivation add to their success or failure in test results.

Technical Guidelines

1. An evaluation plan must establish a schedule of evaluation activities.

   Each evaluation consists of a set of tasks, including data gathering, analysis and reporting. The efficient accomplishment of those tasks require a schedule of activities. Such a schedule also helps in monitoring the progress of evaluation efforts.

2. An evaluation plan should include provisions for providing a timely return of information where appropriate.

The definition of evaluation which was stated earlier indicated that the purpose of the evaluations was to provide a basis for decisions. Therefore, it is imperative that the information is provided before the decision is made. Evaluations that are not timely are not used and thus are probably not worth doing.

3. An evaluation plan must specify role responsibilities in carrying out the plan.

Most evaluations require the cooperation of the evaluation staff and other offices or departments in the school district. Roles and responsibilities of all parties must be delineated so that the evaluation tasks are completed without conflicts.

4. An evaluation plan must identify the reliability and validity of data and data sources.

Not all data in an evaluation are equally reliable and valid. Decision makers need to know the degree of reliability and validity of each datum so that appropriate levels of confidence are given to the data when decisions are made.

5. An evaluation plan should make provision for longitudinal storage and retrieval of data.

The evaluation system which is proposed here is based in part on the idea that the accumulated knowledge of several program evaluations should provide empirical bases for revising district goals and priorities. Such an approach may require secondary analyses or aggregated information from several studies. Therefore data must be stored in such a way that these analyses are possible.

6. An evaluation plan must be cost effective.

The information gained from an evaluation must be worth more than the cost of conducting the evaluation. The identification of program inefficiencies and the payoffs from improved decisions based on evaluative

information cause most evaluations to be cost effective.

Evaluation Requisites

1. An evaluation plan must stimulate a wider use of data and encourage greater reliance of more staff members in planning evaluation activities.

   The increased use of data based decision-making will require that more people identify specific information needs. This situation would then lead to increased participation in planning evaluation activities.

2. An evaluation plan must make provisions for staff development to facilitate the evaluation function.

   This point suggests that staff development is necessary for evaluations to be efficient and successful. The staff development may be necessary for those on the evaluation staff but it is also necessary for persons with whom the evaluators must work while conducting the evaluation study.

3. An evaluation plan should carry teacher and administrator organization support.

   The quality of evaluations depends directly on the quality of the information which is gathered. That information almost always requires the cooperation of others, notably teachers and administrators. Evaluations can be exercised in futility without this support.

4. An evaluation plan must be evaluated to determine the extent to which it has served its purpose.

   Just as the evaluation of programs should lead to the improvement of those programs, so too does the evaluation of evaluations lead to their improvement. Identifying what evaluation procedures are effective will enhance evaluation efforts in the future.

Section II

Evaluation Responsibilities

The evaluation system is designed to be responsive to the information needs of various people and offices in the Toledo Public Schools. The responsibilities of the office are a direct consequence of the requests and changes that others have given to the office and justified as necessary parts of the evaluation effort.

Defining Responsibilities

A systematic effort was undertaken during the fall of 1979 to review the present foci of the evaluation system with an eye toward identifying the unmet needs of various potential user groups. A series of three meetings were held with representatives of present and potential users of evaluation services. At these meetings the present evaluation efforts were discussed and suggestions for improvements were made. More importantly, there were discussions of the information needs of groups of people in the school system so that the high priority unmet needs could be included in a revised evaluation system.

The three planning meetings included:

(1) central office personnel who represented curriculum and instruction, secondary, and elementary education, as well as the business and pupil personnel and administration areas of the school district.

(2) building level personnel as represented by principals and deans.

(3) teachers representating a variety of classroom settings.

The representatives to these meetings were invited through the various professional organizations so that these groups would be aware of the developmental efforts and would have representation in the process from the beginning.

Areas of Evaluation Responsibilities

A large number of present and potential evaluation concerns were discussed

and endorsed as appropriate and important areas of responsibility in the evaluation system. Five major components were determined. More specific information on these areas is included as subsequent sections to this paper. The following paragraphs are general descriptions of these five areas which are depicted in figure 2. More detailed information on these areas is included in the Appendix.

1. Demographic Information

The Evaluation system is charged with the responsibility to gather, synthesize, and report the "longitudinal data base" for the school system. Data on the community, students and parents, and economic and social trends are gathered and used to describe the evolving context within which decisions about the schools must be made. Census data are widely used in this effort but much of the data must be gathered from the primary source, the students and their parents.

There are usually three major levels at which these data are reported, the school system, the individual school building and the community as a whole. These three levels are shown on figure 2.

2. Program Evaluation

The specific educational programs such as the Title I reading program, the English as a Second Language program, or the program for Talented and Gifted Students need to be evaluated. Most of these evaluations are mandated by the provisions of the agreements with the state and federal funding sources.

Since the program evaluation component of figure 2 is the least understood of the five areas and because it may be the most important area in the near future, it is important to describe this area in some detail. An evaluation process that is appropriate for most program evaluations in the Toledo Public Schools appears in Section III, Page 14, Program Evaluation.

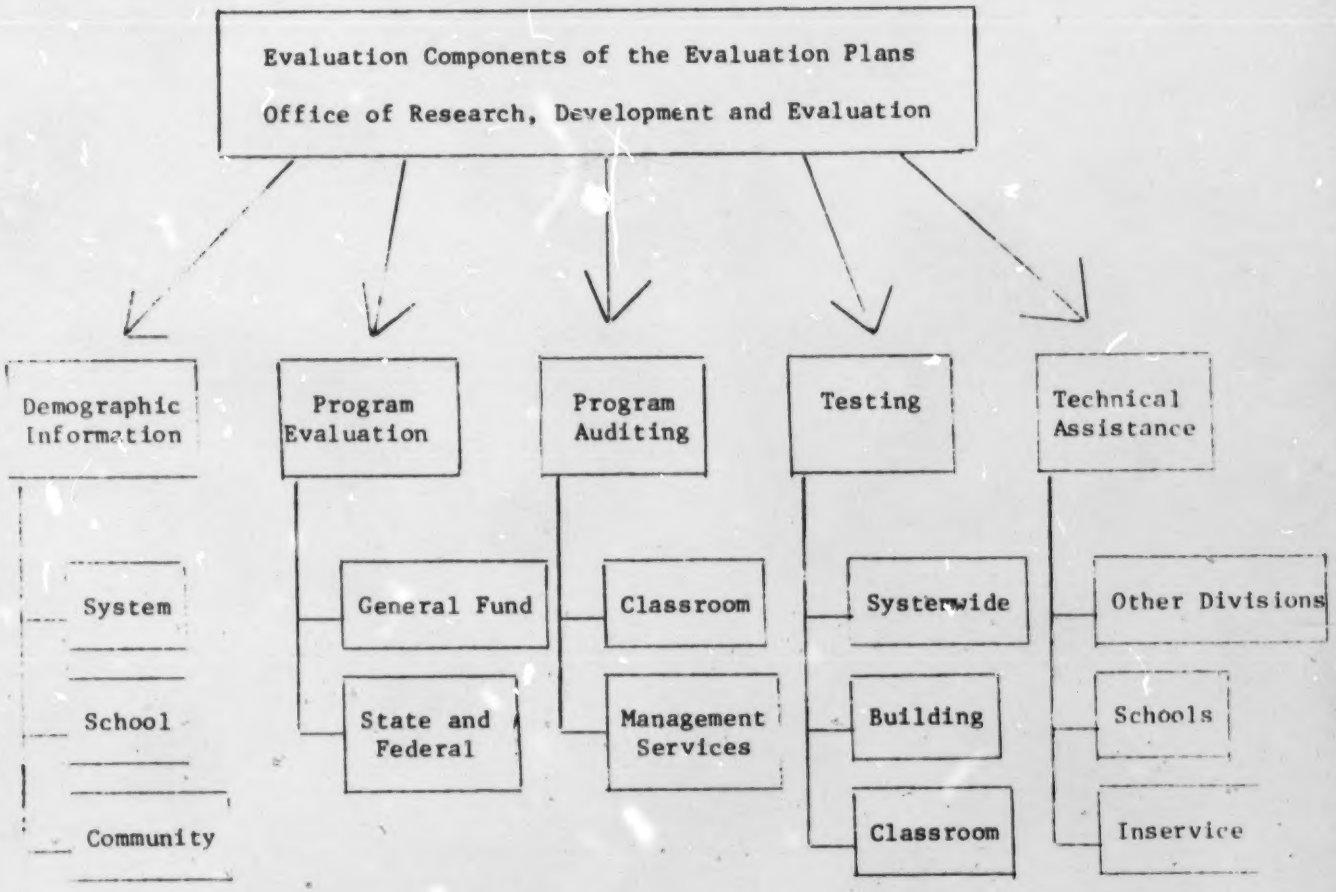


Figure 2. Evaluation Components

3. Program Auditing

There are many funded programs in which the funding sources require a report which gives evidence that the instructional program was implemented in a manner which is consistent with its design, that the staffing and activities which are mandated actually occurred. Similarly there are general fund supported part of the instructional program that will be monitored to see whether their use is consistent with program expectations and/or the inservice training programs and recommended procedures. All of these functions of the evaluation system fall under the program auditing heading.

It is anticipated that much of the program auditing will be done by elementary school principals and high school assistant principals for instruction. An expansion of their job descriptions to include program auditing will require some inservice training and a method for aggregating the data from the separate schools so that it is useful for decision making at the system wide level.

The training programs for these people and the specification of how program auditing fits with other evaluation functions is presently being developed. The fact that many people, teachers as well as administrators, expressed a need for this function suggests that this is an important and integral component in the evaluation system.

4. Testing

It is important to note that the testing program occurs at the system level, the building level and the classroom level. The Office of Research, Development and Evaluation assists or coordinates testing efforts at the building and system level since its resources are limited.

Standardized testing refers primarily to the use of achievement tests in each of the schools. Other standardized tests are given on a system-wide basis and they are described in a separated section.

Testing at the building level might consist of standardized tests or it might consist of other tests, observation systems, or affective scales. It is likely that the selection of measures for building level assessment will rely heavily on the catalogs of instruments which were developed for the Toledo Public Schools by Western Michigan University.

A major component of the achievement assessment system for the Toledo Public Schools will be The Toledo Catalog: A Compilation of Student Behaviors and Administrative Tasks. This component will serve as a resource of assessment instruments and procedures for: (a) a significantly expanded range of student growth areas, which includes behaviors from the cognitive, psychomotor and affective areas, and (b) a newly created listing of management service areas. This effort provides the potential for the Toledo Public Schools to assess almost every aspect of student growth and management effectiveness.

5. Technical Assistance

The Office of Research, Development and Evaluation will provide technical assistance to other divisions in the central office of the Toledo Public Schools and to individual schools in the system. The range of services includes (a) advice on writing measurable objectives so that program outcomes can be evaluated, (b) selecting valid and reliable measurement instruments, (c) advice on appropriate statistical analysis, (d) requests for current information on technical matters such as criterion-referenced testing, and (e) provides more structured, formalized inservice training sessions.

The area of evaluation responsibilities also includes responding to requests from individuals or organizations for information about the district. Many researchers seek the advice of this office or information to complete surveys. These responsibilities require a substantial amount of time and effort.

Interdependency the Five Evaluation Areas

The areas of Demographics, Program Evaluation, Program Auditing, Testing,

and Technical Assistance are not mutually exclusive. A program evaluation may use demographic information and scores from the testing program. A report to the Board of Education or the State Department of Education may require information from all five areas. The components of the evaluation system overlaps and are interdependent. And, because staffing in the office is limited, it is not uncommon for staff members to function across several of these evaluation areas.

Finally, because each of these five components is essential for providing an adequate information base for decisions about the schools, it is essential that each of these areas has a budget commensurate with its time and person requirements. The costs of not having this information greatly exceeds the cost of acquiring it.

Section III

Program Evaluation

The list of evaluation areas that was presented in the previous section includes many instances where specific educational programs such as programs for gifted and talented students or career education need to be evaluated. The following model for program evaluation is meant to be a guide for such evaluations that will probably be appropriate for most situations.

The schematic in Figure 3 indicates five major points where the evaluation office has the primary responsibility in the development of an educational program. A brief description of each of these points is presented so that the processes as well as the roles and responsibilities of various departments and persons are made clear.

1. Generalized Needs Assessment

The identification of problems leads to the development of educational programs to address those problems. Problems are identified through general assessments of need in which accumulated test scores, perceptions of school people and the community, and changes in the world and society are analyzed with regard to identifying discrepancies between what ought to be occurring and what actually is occurring in the schools.

It is the responsibility of the evaluation office to maintain the data on which such an assessment of need is based. Gathering test scores and the perceptions of major groupings of people as well as the shifting demographics of the region must be done on a regular basis. It is also imperative that the data are stored in such a way that they are retrievable and aggregatable so that assessments of needs at various levels can be made.

2. Focused Assessment

Once a problem area is determined, those who are working on the problem

should cooperate with the evaluation office in designing and conducting a focused assessment. For example, if a general assessment suggests that discipline in the junior high schools is a problem, a focused assessment would study the problem in depth for the purpose of providing information about programming options that would address the problem. The focused assessment would investigate the nature and extent of the problem so that corrective programs could be developed.

The focused assessment must be a cooperative effort between those who are charged with developing a program and the evaluation office. The steps in the process should include the following:

a. define target populations - identify the people about whom decisions have to be made

b. select samples - choose representative people from the target population

c. define variables to be assessed - identify skills, knowledge and attitudes to be evaluated

d. develop or select instruments - create or purchase texts, checklists, etc.

e. data gathering - collect the appropriate information through tests, observations, and surveys

f. analysis of data - tabulate and organize the results

g. interpretation of results - compare results to expectations and describe the degree of success

h. reporting to various audiences - presenting written and oral reports of results

3. Formative Evaluation

Formative evaluation deals with identifying the strengths and weaknesses in developing programs for the purpose of program improvement. Information from the evaluation is fed back to the program developers so that they might modify or refine the program based on the most current data.

A trusting relationship with mutual credibility between the evaluators and the program developers is essential in this phase. The usual components

of the formative evaluation are these:

a. specifying Objectives

The evaluation staff assists in ensuring that the objectives of the program are measurable and that appropriate instrumentation is developed.

b. pilot testing

Initial drafts of the program are often submitted to review and comment by experts or relevant groups of people for the purpose of improving the program before it is implemented.

c. field testing

The new program is usually tried out on a small scale before full implementation. The design of the field test, the use of appropriate instruments and comparison groups, the analysis of data and the reporting of the results are primarily the responsibility of the evaluation office.

d. monitoring implementation

When the new program is implemented on a larger scale, it is important to monitor the program to see whether the program is implemented in a manner which is congruent with its design. Information from this activity is fed back to the program developers when implementation is faulty.

4. Summative Evaluation

The summative evaluation of the program refers to an overall estimation of the quality of the education program and its results. Information from the summative evaluation is primarily directed to audiences other than the program developers. The board of education, the community, the superintendent, and funding agents are the usual audiences of summative evaluation information.

Summative evaluations present the benefits which the program produces. These benefits are often described in terms of growth on standardized tests, the accomplishment of specific objectives, the changes in the attitudes of students, and the cost-effectiveness of the program. These conclusions are usually based on comparisons with baseline data, control groups, national or local averages, pre set standards, or completing programs.

A comprehensive summative evaluation will include an assessment of both anticipated and unanticipated program results. The anticipated results are usually stated in the program objectives and are assessed by tests, questionnaires, and observations. Unanticipated or secondary effects are usually obtained from parents, teachers not in the program, and others. Both kinds of information are necessary in assessing the overall effects of the program.

5. Dissemination of Results

An important, but often slighted function of the evaluation office in program evaluation is dissemination of results. It is the responsibility of the evaluation office to work cooperatively with program developers to inform all interested audiences about the results of the program. Such information should go to the profession so that other school districts can share from the experiences of the Toledo Public Schools. Information should also go to parents and the community who fiscally support the programs. Finally, the results should go back to the school, family, the board, administrators, and teachers, so that they are informed about what is happening in their schools.

Summary

Evaluation is an integral part of the change process in the Toledo Public Schools. Specific programs which address educational problems interact --

through the facilitating activities of the evaluation office -- in at least five ways. These interactions have been described in this section.

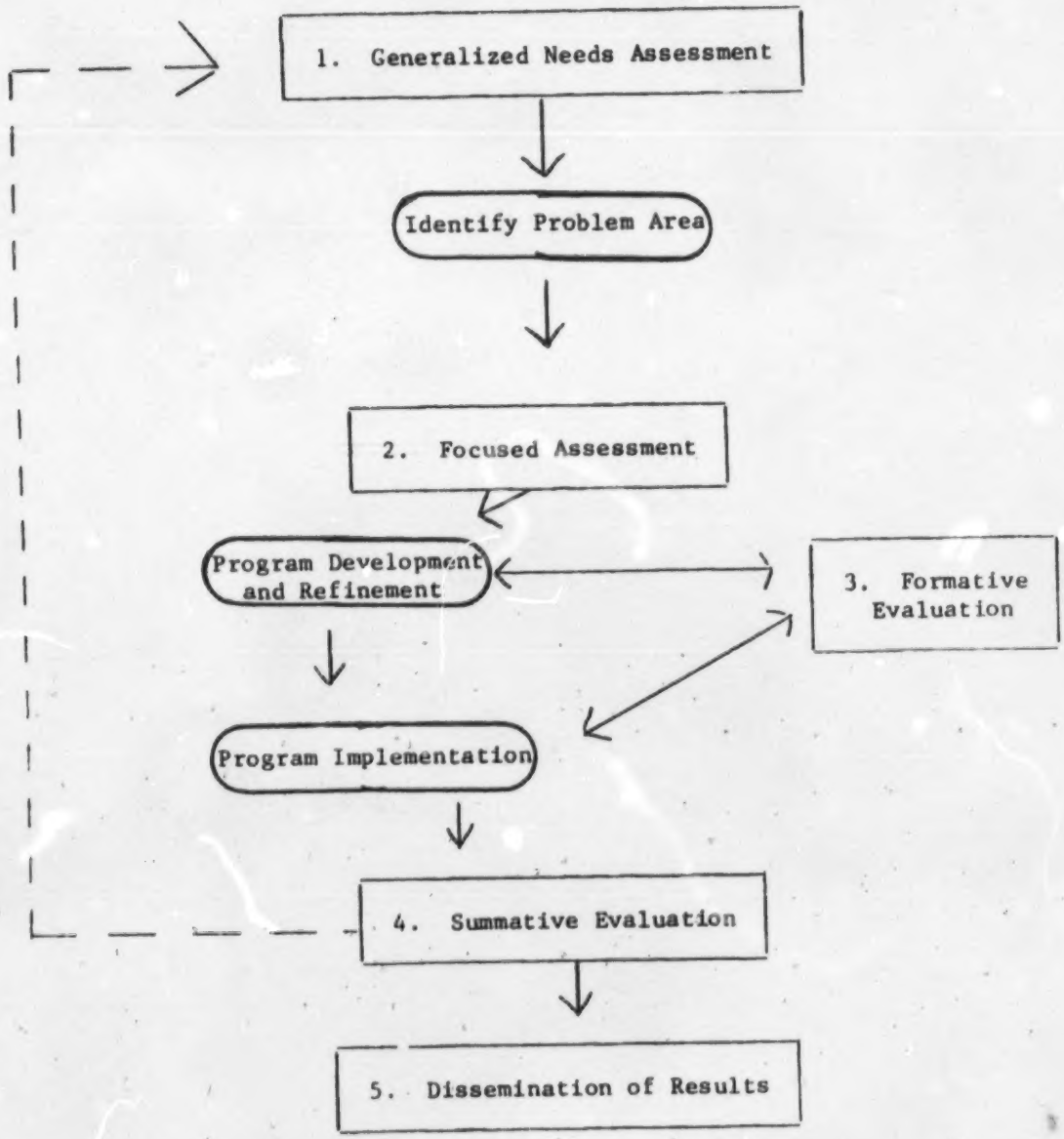


Figure 3. A General Process for Program Evaluation

Section IV

The Recommended Program of Systemwide Standardized Testing

Rationale and Perspective

There has been a need to update the standardized testing progrm in Toledo Public Schools (TPS). The Metropolitan Achievement (MAT) 1970 edition (reading and math) is the only standardized test that has been used systemwide in grades 1-8 for the last eight years. The only systemwide testing in grades 9-12 for the eight year period has been done through the use of Ohio Vocational Interest Survey Test and the General Aptitude Test Battery in the Career Decisions Program at grade 10. No achievement tests have been given systemwide in grades 9-12 for over ten years. There are several factors that dictate the collection of systemwide test data for all grade levels. These are:

- Systemwide comparative standardized test data are needed for curriculum assessment and modification in the Toledo Public Schools. Data from such comparisons are required to assess the outcomes associated with the adoption of new textbook series, new curriculum designs or new instructional programs. Data from such comparisons permit curricular personnel to determine whether students continue to acquire basic skills when innovative instructional programs that might detract from basic skill development impact the curriculum. Data from such comparisons are needed to determine whether the TPS curriculum is remaining current in light of the rapid changes that occur in a complex technological society.
- Standardized test data are needed on individual students to allow teachers, counselors and principals to report to parents the achievement level of their children across the system. Such test data, reported as a group results, give the TPS community one type of information regarding

the success of the educational systems that are presently in operation in TPS.

- Standardized test data on individual students are needed to assist teachers in making appropriate individualized instructional decisions and assignments for the classroom.

- Standardized group test data collection is needed over two or more consecutive years to permit instructional and administrative personnel to detect beneficial or detrimental trends that are occurring in areas of student skill development and achievement.

- Standardized test data on individual students collected system-wide are needed as a key component in the placement of students in such things as honors classes, classes in which highly individualized help is available, some programs funded through state and federal funds and certain junior and high school courses such as algebra.

- Standardized test data need to be collected for use in documenting Needs Assessment sections of project proposals under development to obtain state and federal funds for programs.

The MAT 1970 edition which has been in use in the TPS system will be published and available for only a limited time since it has been replaced by a more recent (1978) edition which contains major changes from the 1970 version. In addition, the MAT 1970 edition appears to have limited validity associated with some aspects of the current TPS curriculum. It was more generally valid in terms of the curriculum nine years ago when the test was adopted by TPS but the curriculum of the schools has changed since that point in time. In December, 1978 a committee of teachers and administrators was organized

for two purposes:

1. To develop recommendations for a Supplemental Testing Program to be implemented in March, 1979 for piloting standardized tests and gathering immediate data for the Model For Quality Education Needs Assessment and for the writing of project proposals.

2. To develop recommendations regarding an on-going comprehensive testing program for TPS to begin with the 1980-81 school year.

The members of the Test Study Committee are:

Dr. Gerald J. Biernacki, Committee Chairman

Mrs. Marlene Davison	Mrs. Aline Stockard
Mrs. Carol Doherty	Mrs. Donna Tuite
Mr. David Jenssen	Mr. Andrew Jordan
Mr. Ron Keaton	Mr. Herman Zielinski

During the course of the study the committee held meetings between December 1, 1978, and February 1, 1980. Curriculum personnel and supervisors of programs which would utilize standardized test results were invited into selected meetings. Their suggestions and input were given consideration during the work performed by the committee.

The Test Study Committee recommended that a Supplemental Testing Program be undertaken (1) to collect sampled test data in selected grade levels in content areas not heretofore available, (2) to explore the test administration merits and demerits of the selected tests and (3) to determine the degree of validity to the TPS curriculum held by each of the tests under consideration for possible adoption into the system-wide testing program. The Supplemental Testing Program was implemented in March, 1979 with students in grades 4, 6, 8 and 10 taking achievement tests. In grades 10 and 12 a

competency test was administered. Competency testing at the twelfth grade was done on a one-time only basis to gather baseline and needs assessment data and was not given consideration as a part of an on-going testing program. The achievement testing associated with the Supplemental Testing program is outlined in Table 1 to include the names of the standardized achievement test batteries which were piloted as well as the subtest areas, grade levels and the numbers of students included.

The committee undertook a complete assessment of the Supplemental Testing Program after its completion and utilized the results in making recommendations regarding an annual, comprehensive achievement testing program for the Toledo Public School System. The assessment of the supplemental test data included an analysis of test contents and an analysis of the administration procedures associated with each instrument. The committee also was concerned with a reasonable balance between the benefits from the tests and the costs in dollars and time.

The Test Study Committee has concluded that a system-wide, balanced, standardized testing program should be adopted. This program would include norm-referenced achievement testing, criterion-referenced achievement assessment and the use of appropriate mini or practice tests. Other types of assessments such as competency tests and attitude inventories are also under consideration by the committee and will be added to the system-wide testing profile if the committee finds they warrant inclusion following future studies and field tests.

In establishing priorities for its work the committee determined that the first and foremost need in the complete assessment profile was in the area of standardized achievement testing. The system-wide achievement test

TABLE 1

Achievement Testing Associated

With The Supplemental Testing Program

MAY 1979

Name of Test	Subtests Administered	Grade Level Tested	Number of Schools*	Number of Students Per School**
Iowa Test of Basic Skills 1978 Edition from Houghton Mifflin	Vocabulary	4	6	30
	Reading Math	6	6	30
	Language Visual Materials Reference Materials	8	3	60
Test of Achievement and Proficiency 1978 Edition from Houghton Mifflin	Reading Math Written Expression Using Sources of Information Social Studies Science	10	3	60
Metropolitan Achievement Test 1978 Edition from Psychological Corporation	Reading	4	6	30
	Math	6	6	30
	Language	8	3	60
	Science (no grade 10) Social Studies (no grade 10)	10	3	60
California Achievement Test 1978 Edition from CTB McGraw-Hill	Reading	4	6	30
	Language	6	6	30
	Math	8	3	60
	Reference Skills	10	3	60
Comprehensive Test of Basic Skills 1975 Edition from CTB McGraw-Hill	Social Studies Science			

*Schools were selected to represent a socio-economic cross-section of the system

**Students were randomly selected from within each building

component of the overall profile is the committee's initial formal recommendation in the development of a complete assessment system for TPS.

Recommendations of the Test Study Committee

The Test Study Committee recommends that the system-wide achievement testing schedule as shown in Table 2 be implemented with the 1980-81 school year with the ITBS and TAP tests. The components of this program are recommended as they appear in the schedule as a result of the following factors:

- The Iowa Test of Basic Skills was selected because:
  - The Iowa Test of Basic Skills (ITBS) 1978 edition,* California Achievement Test (CAT) 1978 edition, and Metropolitan Achievement Test (MAT) 1978 edition showed equally acceptable content validity in terms of the TPS curriculum.
  - The ITBS and CAT offer separate subtest scores in reading and math. The MAT furnishes only a Reading Total and Math Total score.
  - The ITBS publisher (Riverside Publishing which is a division of Houghton Mifflin) provided the most technical support and help throughout the Supplemental Testing Program and therefore appears to be the publisher that would provide the greatest technical support during the period the test was in use in the system.
  - The ITBS publisher has made available a booklet for teachers that correlates each test item with the Houghton Mifflin elementary reading series and a second booklet

*The ITBS is designed for grades kindergarten through eight while the Test of Achievement and Proficiency (TAP) from the same publisher was created for grades 9 through 12.

TABLE 2

Achievement Test Schedule

Recommended By The Test Study Committee

April 1980

Grade Level	Basic or Complete* Battery	Month of Year Testing Should Be Done	Iowa Test of Basic Skills or Test of Achievement & Proficiency
1	Basic	April	ITBS
2	Basic	April	ITBS
3	Basic	April	ITBS
4	Complete	April	ITBS
5	Basic	April	ITBS
6	Complete	December	ITBS
7	Basic	December	ITBS
8	Complete	December	ITBS
9	Basic	December	TAP
10	No achievement testing since career program administers vocationally oriented tests		
11	Basic	December	TAP
12	No testing for this grade level		

*ITBS Basic Battery: Vocabulary, Reading Comprehension, Spelling, Math Concepts, Math Problem Solving, Math Computation

ITBS Complete Battery: Basic Battery + Capitalization, Punctuation, Usage, Visual Materials, Reference Materials

TAP Basic Battery: Reading, Math, Written Expression, Using Sources of Information

correlating the test items with the MacMillan reading series. Both of these reading programs have been adopted and are in use by TPS. Neither the MAT nor the CAT publishers offer material or guides which correlate test data with specific instructional contents for the classroom teacher.

- No major problems were associated with test administration for any of the three achievement tests.
- The cost os systemwide adoption was about equal for all three of the achievement tests.

- The use of the ITBS Basic Battery is distributed across the elementary and junior high grade levels to assess vocabulary, reading comprehension, spelling, math concepts, math problem solving and math computation skills for all students in the system on an annual basis. The Complete Battery is to be used at grade levels 4, 6 and 8 to assess capitalization, punctuation and usage which are key components in the new language arts program. Visual and reference materials subtests from the Complete Battery are to be assessed at these same grade levels to determine student development in these social studies skills. The committee felt that data from these subtests on alternate grade levels would be sufficient to monitor the progress of individual students and to assess these curricular fields while, at the same time, the quantity of testing would be limited.

- The grade levels to be tested were selected for the following reasons:
  - Kindergarten Achievement testing was omitted for the present but will be added in the future at the conclusion of the work being done by the

Kindergarten Assessment Committee.

- Grade 10 was omitted from the achievement testing since students at this grade level are given the Ohio Vocational Interest Survey and the General Aptitude Test Battery in conjunction with the career program.
- Grade 12 was omitted since students at this level have reached the point at which there is too little time left in their programs for supplemental or remedial work if weak skill areas are detected.

- The schedule of testing times (fall, winter or spring) was developed for the following reasons:
  - Thirty of the 58 schools, or 52% of the buildings, containing grades 1-5 must have spring test scores available for these grade levels by the middle of May to determine which students will be eligible for Title I services the following school year. For this reason it is recommended that students throughout the system in grades 1 through 5 be tested in the spring (April).
  - In February, students in grades 6, 7, 8, 9 and 11 must have schedules prepared for their junior or senior high classes the following year. It is therefore recommended that these grade levels be tested in mid-year (December), to provide up-to-date scores in February to determine student placement and to recommend class choices.

The Test Study Committee recommends that, with the adoption of a new achievement test, appropriate inservice be provided to teachers, principals, administrators and counselors. This inservice should center around two factors, test administration procedures and the meaning and use of test results. A plan for these sessions will be created cooperatively by the TPS Test Coordinator, Staff Development and the Riverside Publishing Company. This plan will be recommended to the Test Study Committee for review and approval.

Finally, the Test Study Committee recommends that a review be made of the test score reporting systems presently in use in the schools and of those that are available with the ITBS and TAP tests. Standardized system-wide methods of reporting scores to parents have not been utilized in TPS for two or more years. This area of reporting should receive consideration in the review.

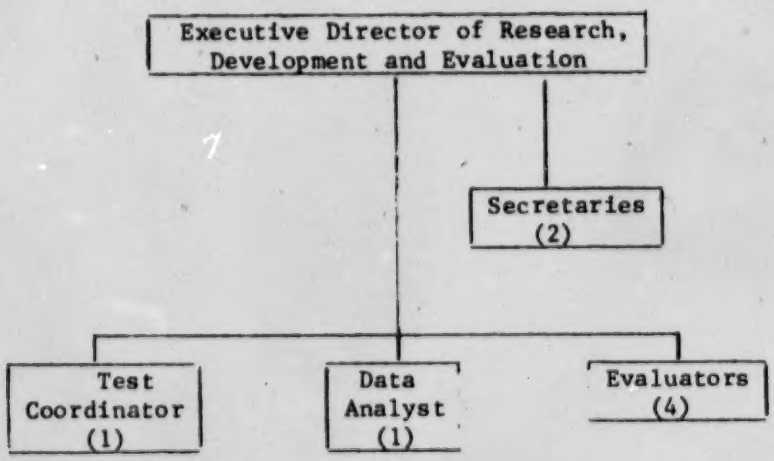
Section V

ORGANIZATION AND DEVELOPMENT

The Office of Research, Development and Evaluation is required to allocate staff to a variety of functions because of the large number of evaluation tasks and the limited evaluation staff. The results of this situation are two-fold: the evaluation staff must serve as generalists in evaluation and second, those who are charged with formative and summative evaluation responsibilities are usually the same people.

Present Staffing

The present staffing strength in the Office of Research, Development and Evaluation is as follows:



These persons presently are responsible for reports on the demographic characteristics of the school system and the community, all of the program evaluations, the program of standardized testing and the provision of technical assistance. Clearly such a limited staff must have expertise in a variety of evaluation areas.

Staffing for Other Evaluation Needs

The increased demand for evaluation services now requires that some of the evaluation functions be carried out at the school building level. The two ways in which this increased evaluative capacity would be realized is through 1) training personnel for program auditing and 2) training personnel as Assessment Coordinators. Both of these roles are described below.

Program Auditors

It is proposed that elementary school principals and high school assistant principals for instruction be trained to perform the function of program auditors. The specific factors that still need to be determined are:

1. Selection criteria, including the specific task description
2. Coordination, the organizational linkage between the school building and the Office of Research, Development and Evaluation
3. Training, the inservice program designed to give these people the necessary skills
4. Link with Decision Based Budgeting, the means for evaluating the cost-effectiveness of the position

The program auditors will be responsible for monitoring instructional programs to attest that the classroom procedures, materials, and activities that are being used are consistent with board policies and the curriculum as designed and adopted by the Toledo Public Schools.

Assessment Coordinators

It is proposed that one person in each school have as a part of his or her responsibilities the function of assessment coordinator. This person would:

A) work with the TFT Building Committee and the Principal in designing the building testing program.

B) assist the Office of Research and Evaluation in the administration of standardized tests and in the interpretation of the results of such tests.

C) assist building personnel in selecting appropriate assessment instruments from the catalogs and other sources and would assist in the interpretation of the results of such testing.

The specific factors that still need to be determined are:

1. Selection criteria, including the specific task description
2. Coordination, the organizational linkage between the school building and the Office of Research and Evaluation
3. Training, the inservice program designed to give these people the necessary skills
4. Link with Decision Based Budgeting, the means for evaluating the cost effectiveness of this position

It is likely that the assessment coordinators would be teachers at the elementary school level and counselors or assistant principals for instruction at the secondary school level.

Training

An important part of the evaluation system is that the evaluation staff remain knowledgeable of the emerging, and recently burgeoning, field of educational evaluation. Recent developments in testing, statistical analysis, evaluation designs, and reporting procedures as well as new guidelines from federal and state funding agents require regular inservice training programs for the evaluation staff. Similarly the close interaction of other offices and departments requires some training of those people so that they understand the demands and procedures of the evaluators. Finally, the expansion of the role of the evaluation office due to the increased demands for information from other offices and audiences may require the training of additional evaluation staff such as the assessment coordinators at the building level.

Other Resources

The very limited evaluation staff cannot possibly respond to all the requests and demands which are brought to them. There are times when certain

evaluation tasks should be carried out but the staff is too overworked or they lack the technical expertise to do the particular job. In these instances it is necessary to tap other resources such as higher education personnel, private contractors, or others who can be contracted on a short-term basis.

The Toledo Public Schools is fortunate to have universities within the region with faculty who can provide such services as well as resources in local industries which can sometimes be used.

Aspects of the Evaluation System that are still in the Developmental Stage

There are two aspects of the evaluation system that still require conceptualization and definition. One of these is the linkage between the evaluation services and fiscal auditing or Decision Based Budgeting. The second aspect is the evaluation of the evaluation system. Each of these concerns is briefly addressed below.

Linkage with Decision-Based Budgeting (DBB)

Some of the evaluation activities are supported by the general fund and some from federal and state funds. It is important that the evaluation services are cost-effective, i.e., that the information which they provide is worth more than the cost of acquiring and reporting that information. Secondly, it is important that the various program evaluations use DBB information in evaluating the worth of instructional programs. Each instructional program has costs associated with it and these need to be considered in any comprehensive evaluation. Often competing educational programs have no demonstrable differences except cost so surely this criterion should be used.

Evaluating the Evaluation System

The evaluation system itself is not above evaluation. Aspects of the system that need to be addressed include 1) the usefulness of the information

that it provides, 2) its credibility, 3) its cost effectiveness, and 4) the extent its products and procedures meet or exceed the Standards for Educational Program Evaluation which have recently been put forward by a joint committee of representatives from various professional educational organizations. Whether this meta-evaluation consists of external professional review, comparative studies, or some other procedures is yet to be determined.

APPENDIX

Included is a sample copy of a detailed information sheet for each of the five component areas of evaluation:

- Demographic Information
- Program Evaluation
- Program Auditing
- Testing
- Technical Assistance

AREA: [illegible]

Planning ___ Operational X

TITLE: Community Demographic Profile

Evaluation Design:	Variables:	Instruments:
	Examples: Family Income per pupil valua- tion Value of houses Percent passing le[illegible]	Census dat[illegible] Tax records Economic data Voting Patterns

Tasks	Responsibility	Schedule
Collect data	Research and Evaluation Joe Rutherford Information Services	On-going
Aggregate data for Annual Report	Research and Evaluation	By November of each ye[illegible]
Develop narrative for Annual Report	Information Services	By Novembe[illegible] o[illegible] each y[illegible]
Distribute Annual Report	Information Services	By November

Evaluation Office Staff: Data Analyst Specialist
Evaluator

Comments:

A community demographic profile is currently available, it will be displayed in the Decision Room, and then updated periodically.

AREA: [illegible]

Planning ___ Operational X

TITLE: Federally Funded Title I Reading and Math

Evaluation Design:	Variables:	Instruments:
CIPP	Reading and Math	Metropolitan Achievement Test
Pre-Post	Perceptions of students, parents and teachers	Student, parent and teacher surveys

Tasks	Responsibility	Schedule
MAT data collection and analysis	Title I teachers pre-test and classroom teachers posttest. Research and Evaluation processes and analyzes data.	First week of October - pre-test April - posttest
Develop, administer and analyze student, parent and teacher surveys	Research and Evaluation	April/May
Aggregate data and write Final Evaluation Report	Research and Evaluation	June/July

Evaluation Office Staff: Program Evaluators
Test Coordinator

Comments:

AREA: Program Auditing

Planning ___ Operational X

TITLE: Screening with Appropriate Assessment Instruments

Evaluation Design:	Variables:	Instruments:
None	Dependent upon program	Dependent upon program

Tasks	Responsibility	Schedule
Review and select tests appropriate to needs of the project	Research and Evaluation and project staff	When and as needed
Conduct and/or audit test administrations	Research and Evaluation and project staff	When and as needed
Analyze test results	Research and Evaluation and project staff	When and as needed
Monitor inclusion of eligible students into program	Research and Evaluation	When and as needed

Evaluation Office Staff:

Program Evaluators
Test Coordinator

Comments:

E.S.E.A. Title I
English Language Development Program

AREA: [illegible]

Planning ___ Operational X

TITLE: Standardized Achievement Test

Evaluation Design:	Variables:	Instruments:
None	Reading Mathematics Study Skills Language Arts Skills	Iowa Test of Basic Skills Test of Achievement and Proficiency

Tasks	Responsibility	Schedule
Coordinate Administration	Research and Evaluation	According to system test schedule
Administer tests	Teachers	According to system test schedule
Coordinate Scoring Services	Research and Evaluation	According to system test schedule
File score reports	Research and Evaluation	According to system test schedule
Provide test results on request for reports and research purposes	Research and Evaluation	According to system test schedule

Evaluation Office Staff:

Test Coordinator

Comments:

Certain printouts from the system-wide testing are provided to schools and teachers. Others are housed in Research and Evaluation for use with curriculum directors, and other administrative personnel.

AREA: Technical Assistance

Planning ___ Operational X

TITLE: Serve As Coordinating Body Between Outside Agency Conducting Research and Toledo Public Schools

Evaluation Design:	Variables:	Instruments:
None	Specified by individual project	Specified by individual project

Tasks	Responsibility	Schedule
Become familiar with goal of study	Research and Evaluation	When-and-as-needed
Make preliminary contacts with Toledo Public School personnel	Research and Evaluation	When-and-as-needed
Audit implementation	Research and Evaluation	When-and-as-needed

Evaluation Office Staff:

Data Analyst
Test Coordinator
Program Evaluators

Comments:

Examples:

National Assessment of Educational Progress
Social Action Research Study
High School and Beyond
Standardized test publishers involved in norming of newly developed tests